A Novel Agent-Based Rumor Spreading Model in Twitter

Emilio Serrano
Universidad Politécnica de Madrid
emilioserra@dit.upm.es

Carlos A. Iglesias
Universidad Politécnica de Madrid
cif@dit.upm.es

Mercedes Garijo
Universidad Politécnica de Madrid
mga@dit.upm.es

ABSTRACT

Viral marketing, marketing techniques that use pre-existing social networks, has experienced a significant encouragement in the last years. In this scope, Twitter is the most studied social network in viral marketing and the rumor spread is a widely researched problem. This paper contributes with a (1) novel agent-based social simulation model for rumors spread in Twitter. This model relies on the hypothesis that (2) when a user is recovered, this user will not influence his or her neighbors in the social network to recover. To support this hypothesis: (3) two Twitter rumor datasets are studied; (4) a baseline model which does not include the hypothesis is revised, reproduced, and implemented; (5) and a number of experiments are conducted comparing the real data with the two models results.

Categories and Subject Descriptors

I.6 [**Simulation and Modeling**]: Model Validation and Analysis; I.2 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*

Keywords

Agent-based Social Simulation, Agent-based Model, Rumor Spreading Model, Information Diffusion Model, Social Networks, Twitter, Big Data

1. INTRODUCTION

Viral marketing, marketing techniques that use pre-existing social networking services, has experienced a significant encouragement over the past few years because a number of reasons. Among others: the low cost of these campaigns; traditional marketing techniques do no longer cause the desired effect; and, people influence each other's decisions considerably [10]. Twitter is the most studied social network in viral marketing. Twitter allows researchers to study global phenomena from a quantitative point of view for the first time in humanity's history [4]. The main reason for this is that, unlike the leading social network Facebook, users' messages in Twitter are public by default.

Rumors are the basis for viral marketing [13] and, therefore, rumors diffusion is a topic widely studied. In this scope, the epidemiological modeling is the hegemonic research line to model the rumor spreading. The standard model in this line is the *SIR* model [9]. In SIR, the population is divided into several classes such as susceptible (S), infected (I), and recovered (R) individuals. Furthermore, these analytical models are usually formulated using differential equations since the transition rates from one class to another are mathematically expressed as derivatives. A different approach is the use of *Agent Based Social Simulation (ABSS)*.

ABSS combines computer simulation and social science by using a simple version of the agent metaphor to specify single components and interactions among them [17][1]. ABSS has become one of the most popular technologies to model and study complex adaptive systems such as emergency management [18], intelligent environments [7], and e-commerce [19]. In the rumor case, ABSS allows researchers to understand how a piece of information spreads on a network and evaluate strategies to control its diffusions; maximizing it in the case of advertisement or minimizing in the case of malicious rumors.

ABSS approaches for rumor spreading, unlike analytical models, allow the exploration of individual-level theories of behavior which can be used to examine larger scale phenomenons [15]. For example, if a single Twitter user gives extensive information for an event while the remaining users post just one tweet (as in Mendoza et al.'s [12] work); ABSS allows this special user to be modeled. The Big Data technologies make the transition from analytical models, which often require overly simplistic assumptions and are difficult to be compared to real-world data [15], to ABSS models a must.

This paper contributes with a (1) novel ABSS model for rumors spread in Twitter. This model relies on the hypothesis that (2) when a user is recovered (the R of the popular SIR model), this user will not influence his or her neighbors in the social network to recover. To support this hypothesis: (3) two Twitter rumor datasets are studied; (4) a baseline ABSS model which does not include the hypothesis is revised, reproduced, and implemented; (5) and a number of experiments are conducted comparing the real data with the two models results.

[1]With some differences, ABSS can also be referred as agent-based models (ABM), multi-agent based simulations (MABS), or social simulation (SocSim).

WWW 2015, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3469-3/15/05.
http://dx.doi.org/10.1145/2740908.2742466.

The paper outline is the following. After revising the related works in ABSS models for rumor spreading in section 2, a baseline model is proposed in section 3. Then, the Twitter data explored is presented in section 4. This section also discusses the concept of recovery and how to model it. Section 5 presents a novel ABSS model for rumor spreading and section 6 experiments with this model and the baseline comparing them with the Twitter data. Finally, section 7 concludes and gives the future works.

2. RELATED WORKS

Although there are a large number of works dealing with rumor spreading by analytical methods, the literature using agent-based modeling is much more reduced. In a recent work, Weng et al. address meme propagations in Twitter. However, getting an accurate spread model is not the paper goal. Tripathy et al. [21] present a study and an evaluation of rumor-like methods for combating the spread of rumors on social networks. They use variants of the independent cascade model [22] for rumor spreading. Liu and Chen [11] build an agent-based rumor spread model using SIR as baseline and implemented in NetLogo [20], a popular ABSS framework. This model is not founded on real data although the authors find out interesting conclusions with regard to the Twitter case using the simulation model. Seo et al. [16] present an over simplistic ABSS. More than the simulation, the contribution rests on the use of this model to evaluate a method to identify rumors and their sources. Yang et al. [23] employ ABSS to analyze the 2013 Associated Press hoax incident. The authors give three profiles for Twitter users (broadcaster, acquaintances, and odd users); probability density functions for each profile; and a study of the effects of removing relevant nodes of the network in the information spread. The Twitter data this model is based on is not given. Gatti et al. [8] address the general information diffusion modeling instead of the rumor spreading. As in other works revised [23], simulation is employed to find users with more impact on the information flow. None of these models consider the hypothesis this paper presents: when a user is "infected" (the I in the SIR model) and then recovered (R); these users may not affect their neighbors' recovery.

3. BASELINE APPROACH

This section revises Tripathy et al.'s approach [21] for modeling the rumor spreading in Twitter. This model is based on the cascade model [22] and is one of the earliest ABSS rumor spread models proposed for the Twitter case. Besides, as in this paper, the authors criticize epidemic spread models such as SIS and SIR because, among others, anti-rumors can be spread from person to person unlike vaccines for viruses which can only be administered to individuals.

In the revised model, Agents are Twitter users with a state property which can be: *neutral* (initial state); *infected* (believe the rumor); *vaccinated* (believe the anti-rumor before being infected); or, *cured* (believe the anti-rumor after being infected). The basic agents' behavior involves: (1) initializing a number of infected users; (2) each infected user at time t tries to infect each of its uninfected neighbors with a given probability (*propInfect*); (3) after a given delay (*timeLag*), a random infected node starts an anti-rumor spreading to its neighbors, trying to cure or vaccinate them with a probability (*probAcceptDeny*) each time step t; and, (4) cured and vaccinated users also try to cure or vaccinate their neighbors with a probability (*probAcceptDeny*) each time step t .

Concerning the environment, a synthetic network is employed. *Barabási Albert* (BA) scale-free networks are the most popular option when modeling social networks [11]. Although the scale-free nature of a large number of networks is still debated by the scientific community, social networks such as Twitter are widely claimed to be scale-free. In a nutshell, the creation of these networks is undertaken under the assumption that the probability a user u_1 connects to another u_2 depends on the number of connections that u_2 already has. To give more information to reproduce this environment, this paper experiments with 1K nodes and a maximum of 10 links initially added per new node. More specifically, the Barabási-Albert preferential attachment graph generator of the graph stream project has been employed[2].

A time step of an hour is assumed, and the output is the number of users endorsing the rumor (with infected as state) and the number of users denying it (with vaccinated or cured as states). With the details given, the input parameters in the baseline approach are the following: random seed, number of users, maximum links per node (for the BA network construction), initially infected users, probability of infecting, probability of accepting a denial, and time lag.

4. TWITTER DATASET AND THE RECOVERY CONCEPT

The baseline model, reproduced and implemented in this paper, and the proposed spread model have been validated using two rumor datasets introduced by Qazvinian et al. [14]. The first dataset called "obama" includes tweets which spread misinformation that president Obama is Muslim. The second dataset called "palin" deals with Sarah Palin divorce rumors. Although the cited work includes other rumors, these were the topics with more tweets retrieved: 4975 for the obama dataset and 4423 for the palin dataset. Hence, they were the most useful for simulation purposes.

Qazvinian et al. [14] not only retrieved tweets based on regular expressions, `Obama & (muslim|islam)` for the obama dataset and `palin & divorce` for the palin dataset, but also annotated manually these tweets. The possible labels for the dataset are: *endorsers* (it spreads the rumor), *denies* the user refutes the rumor), *questions* (the user questions the rumor credibility), *neutral* (the tweet is about the rumor without endorsing or denying it), *unrelated* (the tweet is not about the rumor), and *undetermined* (when the annotator is undetermined). The mere existence of the "undetermined" label, used when a human annotator cannot decide, illustrates the challenging problem of automatically detecting if a tweet is a rumor or not which, although is out of the scope of this paper, is a hot research topic.

The explained datasets were provided in different formats. In this paper, their tweets have been: (1) retrieved again from the id when available by using the Twitter REST API [2] (obama case); (2) extended by retrieving retweets of the original tweets (obama case); (3) anonymized for their distribution obeying the Twitter terms of use [3] (palin case);

[2]GraphStream project: http://graphstream-project.org/

and, made available at this paper additional material website under a creative commons license [1].

After studying Twitter data of these and other datasets, the authors found out that the "recovery" concept which most popular rumor spread model relies on is complex of being validated. The main reason is that when retrieving tweets about rumors (or anti-rumors) in a specific topic, all the information for most of the users usually comes from just one tweet which says if the user is endorsing or denying the rumor. Therefore, even if the user has been "cured" of the rumor, there is not empirical evidence of it.

5. A NOVEL SPREAD MODEL

The new model introduced in this paper modifies the baseline to include the idea of users who may know that a rumor is false but who do not spread anti-rumors, i.e. tweets denying the rumors. As explained, the main reason for introducing this idea is that rumor information for a specific user is typically limited to just one tweet. Furthermore, psychologically, the infected users who make a mistake may not be as enthusiastic as the baseline model assumes about spreading their faults with anti-rumors.

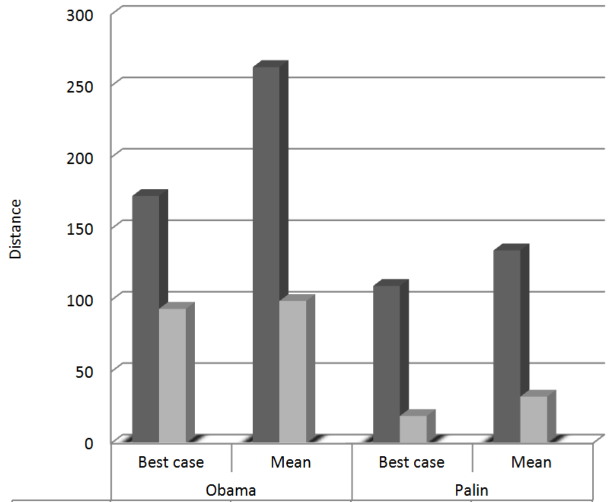
As a result, only vaccinated users (the ones who have not been previously infected) are allowed to spread anti-rumors. Another idea included in the new model is that, independently of any *time lag*, a neutral node which has an infected neighbor can become a vaccinated user. This models the idea of a user who knew from any external information that the rumor was false. With this in mind, a *probability of making a denier* is included, i.e. turning a neutral user into a vaccinated user when spreading a rumor.

Thus, the agents' behavior in the new model is: (1) initializing a number of infected users; (2) each infected user at time t tries to infect each of its uninfected neighbors with a given probability (*propInfect*); (3) instead of infecting them, these neighbors may become vaccinated if they were neutral with a probability (*propMakeDenier*); (4) vaccinated users (but not cured users) attempt to cure or vaccinate their neighbors with a probability (*probAcceptDeny*) each time step t.

With the details given, the input parameters in the new model are the same as in baseline except the time lag which is replaced with *propMakeDenier*: random seed, number of users, maximum links per node (for the BA network construction), initially infected users, probability of infecting, probability of accepting a denial, and probability of making a denier.

6. EXPERIMENTAL RESULTS

This section further supports the hypothesis that is more realistic to consider that users who have spread a rumor will not spread anti-rumors in Twitter. For that purpose, the baseline approach and the novel spread model have been implemented to compare their results with the two Twitter datasets explained in section 4. More specifically, the number of users endorsing and denying a rumor in the simulation is compared to the number of these users in the real data. Thus, the following *distance metric* is used to validate the realism of the simulations: $d(endorsers, sim, dataset) + d(deniers, sim, dataset)$, where d calculates the Euclidean distance between the number of users (nu) of a specific



	Obama		Palin	
	Best case	Mean	Best case	Mean
Baseline	172.535	262.696	109.572	134.463
Proposed	93.51	99.022	18.546	32.095
Reduction	45.80%	62.31%	83.07%	76.13%

Figure 1: Experimental results.

type (endorser or denier) in the simulation and the *dataset* (obama or palin) for the days considered (*nDays*).

The number of endorsers and deniers users is calculated differently for datasets, the baseline model, and the proposed model. In the datasets, a user is counted as endorser or denier if his or her last tweet was labelled as *endorsers* or *denies*, respectively. In the baseline approach, *infected* users count as endorsers and, *vaccinated* and *cured* as deniers. In the proposed model, cured agents are counted as users endorsing the rumor along with the infected agents; and only vaccinated agents are counted as users denying the rumor.

The implementation uses, among others, the MASON Multiagent Simulation Toolkit[3] and the GraphStream project. The employed parameters combinations give over 173K experiments for the baseline and over 170K experiments for the proposed model.

Figure 1 shows the main results. The figure shows the distances for the obama and palin datasets: (1) in the best case achieved by the models; (2) and, in the best mean of distances considering all random seeds for a parameter values set[4]. These results show that the proposed model achieves reductions of distance between 45.80%, obama best case, and 83.07%, palin best case. Extended experiments results are available online in the additional material web [1].

7. CONCLUSIONS AND FUTURE WORKS

The epidemiological modeling is the hegemonic approach to model rumor spreading. This paper challenges that approach by assuming that users who realize that have spread a false rumor typically: (1) will not spread anti-rumors, or (2) there will not be empirical evidence of the retraction. Therefore, the recovered users will not affect the recovery of their neighbors. The exploratory data analysis of two Twitter rumor datasets about Obama and Palin supports this

[3]Mason: `http://cs.gmu.edu/~eclab/projects/mason/`

[4]Standard deviations are also provided in the extended experiments [1].

hypothesis. Concretely, most of the users information comes from a single tweet which allows researchers to consider them rumor endorsers or deniers, making the "recovery" perception unavailable. A novel agent-based rumor spread model considering the explained hypothesis is introduced in this paper. The proposed model is compared to one of the earliest agent-based social simulation (ABSS) rumor spread models for the Twitter case. Experimental results show that the novel model is able to reduce between 45% and 83% the distance with the two Twitter datasets studied. Both the Twitter datasets and the complete experimental data are available online [1].

Our main future work is the integration of the presented model with Big Data technologies to: explore how to automatically model Twitter users; compare these models with historic and online data generated by these users; to adjust the model parameters with the online data generated; and to suggest actions for maximizing or minimizing the information spread.

Acknowledgments

This research work is supported by the Spanish Ministry of Economy and Competitiveness under the R&D project CALISTA (TEC2012-32457); by the Spanish Ministry of Industry, Energy and Tourism under the R&D project BigMarket (TSI-100102-2013-80); and, by the Autonomous Region of Madrid through the program MOSI-AGIL-CM (grant P2013/ICE-3019, co-funded by EU Structural Funds FSE and FEDER).

8. REFERENCES

[1] Paper additional material, datasets and extended experiments. `http://goo.gl/ImJCKt`.

[2] Twitter REST API documentation website. `https://dev.twitter.com/overview/documentation`.

[3] Twitter terms of use website. `https://dev.twitter.com/overview/terms/policy`.

[4] M. De Domenico, A. Lima, P. Mougel, and M. Musolesi. The Anatomy of a Scientific Rumor. *Scientific Reports*, 3, Oct. 2013.

[5] J. D. Farmer and D. Foley. The economy needs agent-based modelling. *Nature*, 460(7256):685–686, Aug. 2009.

[6] F. Flentge, D. Polani, and T. Uthmann. Modelling the emergence of possession norms using memes. *J. Artificial Societies and Social Simulation*, (4).

[7] T. Garcia-Valverde, F. Campuzano, E. Serrano, A. Villa, and J. A. Botia. Simulation of human behaviours for the validation of ambient intelligence services: A methodological approach. *Journal of Ambient Intelligence and Smart Environments*, 4(3):163–181, Aug. 2012.

[8] M. A. d. C. Gatti, A. P. Appel, C. N. dos Santos, C. S. Pinhanez, P. R. Cavalin, and S. B. Neto. A simulation-based approach to analyze the information diffusion in microblogging online social network. In *Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World*, WSC '13, pages 1685–1696, Piscataway, NJ, USA, 2013. IEEE Press.

[9] H. W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42:599–653, 2000.

[10] J. Kostka, Y. A. Oswald, and R. Wattenhofer. Word of Mouth: Rumor Dissemination in Social Networks Structural Information and Communication Complexity. In A. A. Shvartsman and P. Felber, editors, *Structural Information and Communication Complexity*, volume 5058 of *Lecture Notes in Computer Science*, chapter 16, pages 185–196. Springer Berlin / Heidelberg, Berlin, Heidelberg, 2008.

[11] D. Liu and X. Chen. Rumor propagation in online social networks like twitter – a simulation study. In *Proceedings of the 2011 Third International Conference on Multimedia Information Networking and Security*, MINES '11, pages 278–282, Washington, DC, USA, 2011. IEEE Computer Society.

[12] M. Mendoza, B. Poblete, and C. Castillo. Twitter under crisis: Can we trust what we rt? In *Proceedings of the First Workshop on Social Media Analytics*, SOMA '10, pages 71–79, New York, NY, USA, 2010. ACM.

[13] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili. Theory of rumour spreading in complex social networks. *Physica A: Statistical Mechanics and its Applications*, 374(1):457 – 470, 2007.

[14] V. Qazvinian, E. Rosengren, D. R. Radev, and Q. Mei. Rumor has it: Identifying misinformation in microblogs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '11, pages 1589–1599, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.

[15] W. Rand and R. T. Rust. Agent-based modeling in marketing: Guidelines for rigor. *International Journal of Research in Marketing*, 28(3):181–193, Sept. 2011.

[16] E. Seo, P. Mohapatra, and T. Abdelzaher. Identifying rumors and their sources in social networks. 2012.

[17] E. Serrano, P. Moncada, M. Garijo, and C. A. Iglesias. Evaluating social choice techniques into intelligent environments by agent based social simulation. *Information Sciences*, 286(0):102–124, 2014.

[18] E. Serrano, G. Poveda, and M. Garijo. Towards a holistic framework for the evaluation of emergency plans in indoor environments. *Sensors*, 14(3):4513–4535, 2014.

[19] E. Serrano, M. Rovatsos, and J. A. Botía. Data mining agent conversations: A qualitative approach to multiagent systems analysis. *Information Sciences*, 230(0):132 – 146, 2013.

[20] S. Tisue and U. Wilensky. NetLogo: A Simple Environment for Modeling Complexity. 2004.

[21] R. M. Tripathy, A. Bagchi, and S. Mehta. A study of rumor control strategies on social networks. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 1817–1820, New York, NY, USA, 2010. ACM.

[22] L. Weng, F. Menczer, and Y.-Y. Ahn. Virality prediction and community structure in social networks. *Scientific Reports*, 3, Aug. 2013.

[23] S. Y. Yang, A. Liu, and S. Y. K. Mo. Twitter financial community modeling using agent based simulation. SSRN scholarly paper, Rochester, NY. IEEE Computational Intelligence in Financial Engineering and Economics, London, 2013.